Chapter 1

Random Processes

1.1 Introduction

The probability theory and stochastic processes become a basic tool designing of digital communication systems including:

- Modeling the sources that generate the information.
- Digitization of the source output.
- Characterizing of various communication channels.
- Designing of the receiver structure.
- Evaluation of the performance of the whole system.

1.2 Concepts of Probability

The theory of probability deals with averages of mass phenomena occurring sequentially or simultaneously electron emission, telephone calls, radar detection, quality control, system failure, noise, birth and death rates, and queuing theory, among many others. The purpose of the probability theory is to describe and predict such averages in terms of probabilities of events.

If an experiment is performed n times and the event A occurs N_a times, then with a high degree of certainty, the relative frequency N_a/N of the occurrence of A is close to the probability of event A. So the **probability of an event *A*** is determined by:

$$P(A) = \frac{N_a}{N} \tag{1.1}$$

Where N is the number of all possible outcomes and N_a is the number of outcomes that are favorable to the event A.

The **union *A+B*** of two events A and B is the event that occurs when A or B or both occur. The **intersection**, or equivalently the joint, ***A.B*** of the events A and B is the event that occurs when both events A and B occur. Generally, if the events A and B interleaved each other, those union and joint probabilities are interrelated by the following expression:

$$P(A+B) = P(A) + P(B) - P(A.B) \tag{1.2}$$

The events A and B are **mutually exclusive** if the occurrence of one of them precludes the occurrence of the other. In such case they are called **independent** so that the union can be given simply by:

$$P(A+B) = P(A) + P(B) \tag{1.3}$$

Now the **conditional probability** can be defined as follows. If the outcomes common to the two events *A* and *B* form their joint *A.B*, then the conditional probability of event *A*, given event *B*, is defined as:

$$P(A/B) = \frac{P(A.B)}{P(B)} \tag{1.4}$$

One can think of the conditional probability, as representing the likelihood of event *A's* having occurred when it is known that event *B* has occurred. However if the events *A* and *B* are independent, then there is no joint occurrence, so that:

$$P(A/B) = P(A) \tag{1.5}$$

Hence, the **joint probability** of two independent events equals to the product of their individual probabilities as indicated in the following:

$$P(A.B) = P(A).P(B) \tag{1.6}$$

Finally, we consider a very practical problem. It involves any experiment for which there are only two possible outcomes on any trial. Hitting or missing the target in artillery, passing or failing an exam, receiving a `0' or a `1' in a digital bit stream transmission, and occurrence or non-occurrence of any event are just a few examples.

Assume p is the probability of occurrence of an event *A*. So, the probability of non-occurrence of such event will be termed as q whereas $q=1-p$.

After repeating the experiment *N* times, the probability that *A* is observed exactly *k* times out of *N* trials, is given by the ***Binomial*** distribution. This may be represented as follows:

$$P(\,A\ occurs\ exactly\ k\ times\,) = \frac{N!}{k!\,(N-k)!}\,p^{k}(1-p)^{N-k} \qquad (1.7)$$

It is important to note that this Binomial distribution dealing with combined experiments or repeated trials of a single experiment.

That is to select *n* from *N* simultaneously in 1-trial or to select *n* in successive repeated *n*-trials.

1.3 Random Variables

A random variable is a number assigned to every outcome of an experiment. This number could be the voltage of a random source, the phase of a random signal, the power of a received signal, or any other numerical quantity that of interest in the performance of the experiment.

The concepts concerning random variables are too long. Only the main topics used in this chapter are summarized to make it useful, readable, and simple without need to further investigation on various literatures.

A random variable *X* is characterized by three basic functions that allow for ready evaluation of any probabilistic question about the random variable. The most fundamental function is the **cumulative distribution function *cdf*** (or simply the distribution function) defined by the following expression:

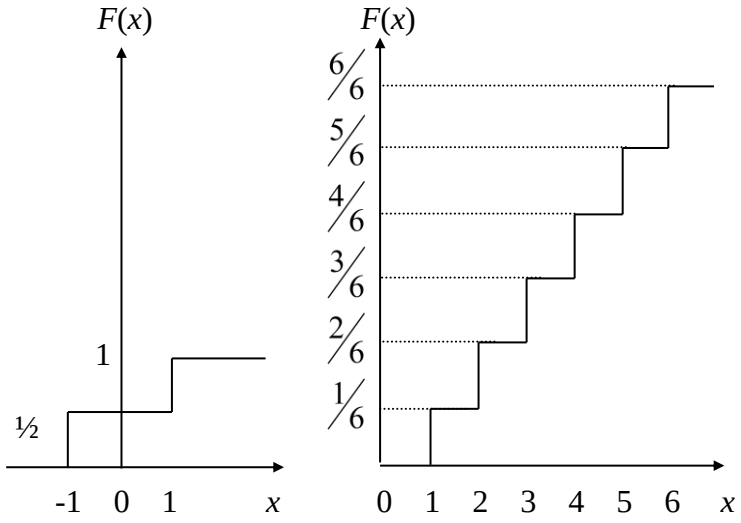
$$F_x(x) = p(X \le x) \qquad (1.8)$$

That is the probability of the event, "***The random variable X takes on a value equal to or less than x***", in a trail. The cumulative distribution function *cdf* is characterized by the following properties:

- $F_x(-\infty)=0$ (1.9)
- $F_x(\infty)=1$ (1.10)
- $0 \le F_x(x) \le 1$ (1.11)
- $F_x(x_1) \le F_x(x_2)$ *if* $x_1 \le x_2$ (1.12)
- $p(x_1 < X \le x_2) = F_x(x_2) - F_x(x_1)$ (1.13)
- $F_x(x^+) = F_x(x)$ (1.14)

For example, the discrete random variable generated by flipping a fair coin has the ***cdf*** shown in Fig.1.1.a. There are two jumps in $F(x)$, one at x = -1 and one at x = 1.

Similarly, the random variable generated by tossing a fair die has the cdf as shown in Fig.1.1.b with 6 jumps, one at each of the points x = 1, 2, …, 6.



(a) Tossing a Coin (b) Tossing a Die

Fig.1.1 Examples of Discrete CDF

The second important function for the random variables statistics is called the **probability density function *pdf***. The probability density function *pdf* of a random variable *X* is the derivative of the cumulative distribution function *cdf* defined as in the following:

$$f_x(x) = \frac{d}{dx} F_x(x) \tag{1.15}$$

Since the density function, *pdf*, is the derivative of the cumulative function, *cdf*, also the cumulative function, *cdf*, can be given in terms of the density function, *pdf*, by integration:

$$F_x(x) = \int_{-\infty}^{x} f_x(y)\, dy \tag{1.16}$$

The probability density function *pdf* is also characterized by the following important properties:

- $f_x(x) > 0 \quad for\ all\ x$ (1.17)
- $\int_{-\infty}^{\infty} f_x(x)\, dx = 1$ (1.18)
- $F_x(x) = \int_{-\infty}^{x} f_x(y)\, dy$ (1.19)
- $P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_x(x)\, dx$ (1.20)

Finally, the **mass function** is usually concerned with discrete random variables.

Meanwhile, a mass function of a random variable denoted by $p_x(x)$ is then defined as in the following:

$$p_x(x) = p(X = x) \tag{1.21}$$

The probability density function given in equation 1.15 is for continuous random variables. If a random variable is discrete taking the values x_i with probability mass functions given by:

$$p_i = p(X = x_i) \tag{1.22}$$

Then, the probability density function can be evaluated in terms of the discrete mass functions as follows:

$$f_x(x) = \sum_i p_i \ \delta(x - x_i) \tag{1.23}$$

It may be necessary to identify the outcomes of an experiment by two (or more) random variables. These random variables may be or may not be independent of one another. For two random variables *X* and *Y*, the probability that $x \le X \le x + dx$ while at the same time $y \le Y \le y + dy$ is given by:

$$P(x \le X \le x + dx, y \le Y \le y + dy) = f_{xy}(x,y)\, dx\, dy \tag{1.24}$$

Extending this for a finite interval leads to:

$$P(x_1 \le X \le x_2, y_1 \le Y \le y_2) = \int_{y_1}^{y_2} \int_{x_1}^{x_2} f_{xy}(x,y)\, dx\, dy \tag{1.25}$$

The cumulative distribution function *cdf* of two random variables:

$$F_{xy}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{xy}(x,y)\, dx\, dy \tag{1.26}$$

Concerning the cumulative distribution function *cdf* for some value of *x* quite independently of *y*, is given by:

$$F_x(x) = P(X \le x, -\infty \le y \le \infty) = \int_{-\infty}^{\infty} \int_{-\infty}^{x} f_{xy}(x,y)\, dx\, dy \tag{1.27}$$

Henceforth, the probability density function *pdf* will be given as follows:

$$f_x(x) = \frac{d}{dx} F_x(x) = \int_{-\infty}^{\infty} f_{xy}(x,y)\, dy \tag{1.28}$$

If the random variables X and Y are independent, the following expressions results:

$$P(x \le X \le x+dx, y \le Y \le y+dy) = \left[f_x(x)\, dx\right]\left[f_y(y)\, dy\right] \tag{1.29}$$

$$P(x_1 \le X \le x_2, y_1 \le Y \le y_2) = \left[\int_{x_1}^{x_2} f_x(x)\, dx\right]\left[\int_{y_1}^{y_2} f_y(y)\, dy\right] \tag{1.30}$$

$$f_{xy}(x,y) = f_x(x)\, f_y(y) \tag{1.31}$$

1.4 Statistical Averages of Random Variables

If the possible numerical values of the random variable X are x_1, x_2, x_3,…, with probability of occurrence $P(x_1)$, $P(x_2)$, $P(x_3)$,…etc. As the number of measurements N becomes very large, the outcome $X=x_1$ is expected to occur $N\,P(x_1)$ so that:

$$x_1 P(x_1) N + x_2 P(x_2) N + \ldots = N \sum_i x_n P(x_n) \tag{1.32}$$

The mean or average value of all these measurements is called the **average value** or the **expectation** of the random variable X, and is calculated by dividing the above sum by N:

$$\overline{X} = E[X] = m = \sum_i x_i P(x_i) \tag{1.33}$$

The average for a continuous random variable is given as follows:

$$\overline{X} = E[X] = m = \int_{-\infty}^{\infty} x\ f(x)\,dx \tag{1.34}$$

Furthermore, the average value of a function *g(X)* of the random variable *X* is given by:

$$\overline{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)\ f(x)\,dx \tag{1.35}$$

Moreover, if the random variable *X* is raised to a power *n*, the average value of X^n is referred to as the ***n<u>th</u>*** **moment** of the random variable *X* and will be given by:

$$\overline{X^n} = E[X^n] = \int_{-\infty}^{\infty} x^n f(x)\,dx \tag{1.36}$$

The **variance** σ^2 of a random variable is a measure of the width of the probability density function. It is equivalent to the average value of the second moment $(X\text{-}m)^2$ as follows:

$$\sigma^2 = E[(X-m)^2] = \int_{-\infty}^{\infty} (x-m)^2\ f(x)\,dx \tag{1.37}$$

$$\begin{aligned}\sigma^2 &= E[(X-m)^2] = E[X^2 - 2mx + m^2] \\ &= E[X^2] - 2mE[X] + m^2 \\ &= E[X^2] - 2m^2 + m^2 = E[X^2] - m^2 \\ &= E[X^2] \qquad if \quad m = 0\end{aligned} \tag{1.38}$$

1.5 Gaussian Probability Density

The Gaussian probability density function is of a greatest importance in communications because many natural events are characterized by random variables with a Gaussian density such as the thermal noise. The Gaussian probability density function is defined as in the following:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2} \tag{1.39}$$

Where m and σ^2 are the mean and variance of the Gaussian function given as follows:

$$m = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2} dx \tag{1.40}$$

$$\sigma^2 = \int_{-\infty}^{\infty} \frac{(x-m)^2}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2} dx \tag{1.41}$$

The area under the Gaussian probability density function is unity:

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \tag{1.42}$$

1.6 Error Function

The cumulative distribution function *cdf* corresponding to the Gaussian probability density *f(x)* with zero average (m=0) is:

$$F(x) = P(X \le x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} dx \tag{1.43}$$

This integral is not easy and is readily available in mathematical tables and is termed as the error function. The **error function** of *u*, written as *erf u* is defined as:

$$erf\ u = \frac{2}{\sqrt{\pi}} \int_0^u e^{-u^2} du$$
$$erf(0) = 0 \qquad \& \qquad erf(\infty) = 1 \tag{1.44}$$

The complementary error function, written as *erfc u*, is defined as:

$$erfc\,u = 1 - erf\ u = \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} e^{-u^2}\,du \tag{1.45}$$

So, the cumulative distribution in equation 1.42 may be expressed in terms of the error function as follows:

$$F(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-x^2/2\sigma^2}\,dx - \int_{x}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-x^2/2\sigma^2}\,dx$$

The first term is equal to unity. If $u = x/\sqrt{2}\sigma$ then:

$$\begin{aligned} F(x) &= 1 - \int_{x/\sqrt{2}\sigma}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-u^2} \sqrt{2}\sigma\,du \\ &= 1 - \frac{1}{2}\left(\int_{x/\sqrt{2}\sigma}^{\infty} \frac{2}{\sqrt{\pi}}\, e^{-u^2}\,du \right) = 1 - erfc\left(\frac{x}{\sqrt{2}\sigma} \right) \end{aligned} \tag{1.46}$$

The *erfc u* is only readily available for positive *u*. For $x < 0$ then:

$$\begin{aligned} F(x) &= F(-|x|) = \int_{-\infty}^{-|x|} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-x^2/2\sigma^2}\,dx \\ &= \int_{-\infty}^{-|x|/\sqrt{2}\sigma} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-u^2} \sqrt{2}\sigma\,du = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{-|x|/\sqrt{2}\sigma} e^{-u^2}\,du \end{aligned} \tag{1.47}$$

Letting $\xi = -u$ yields:

$$F(x) = \frac{1}{2}\left[\frac{2}{\sqrt{\pi}} \int_{|x|/\sqrt{2}\sigma}^{\infty} e^{-u^2}\,du \right] = \frac{1}{2}\, erfc\left(\frac{|x|}{\sqrt{2}\sigma} \right) \tag{1.48}$$

Now, given a random variable *X* and its associated density function $f_X(x)$, what is the density function of the random variable $Y = g(x)$ where $g(x)$ is some function of *X*?.

To find the probability density $f_y(y)$ for a specific y, we solve the equation $Y=g(x)$. If it has x_n real roots and $g'(x)$ is the derivative of $g(x)$, then the density function will be given by the following expression:

$$f_y(y)=\frac{f_x(x_1)}{g'(x_1)}+\frac{f_x(x_2)}{g'(x_2)}+...+\frac{f_x(x_n)}{g'(x_n)} \tag{1.49}$$

As a summary it is convenient to indicate the following important cases that concerns the density function of Y in terms of the density function of X.

- If $g(x)$ is a linear function of X, such as $Y = a\ X + b$, the derivative will be $g'(x) = a$ and hence:

$$f_y(y)=\frac{1}{a}f_x\left(\frac{y-b}{a}\right) \tag{1.50}$$

- If $g(x)$ is the inverse so that $Y = 1/X$, the derivative will be $g'(x) = -1/x^2$. Then the density function is given as in the following expression:

$$f_y(y)=\frac{1}{y^2}f_x\left(\frac{1}{y}\right) \tag{1.51}$$

- If $g(x)$ is the squarer of X (i.e., $Y = aX^2$), the derivative $g'(x)$ is given by $2ax$, where a is greater than 0. The density of Y:

$$f_y(y)=\frac{1}{2a\sqrt{y/a}}\left[f_x\left(\sqrt{\frac{1}{y}}\right)-f_x\left(-\sqrt{\frac{1}{y}}\right)\right] \quad ,y>0 \tag{1.52}$$

- If $g(x)$ is in the sinusoidal form $Y= a\ sin(X+\phi)$ of n solutions $f_x(x_n)$ with derivatives:

$$g'(x_n) = \sqrt{a^2 + y^2} \tag{1.53}$$

Hence, the density function of Y is given by:

$$f_y(y) = \frac{1}{\sqrt{a^2 - y^2}} \sum_{n=-\infty}^{\infty} f_x(x_n) \qquad , y < a \tag{1.54}$$

1.7 The Central Limit Theorem

The probability density of a sum of N independent random variables tends to approach a Gaussian density as the number N increases. The mean and variance of this Gaussian density are respectively the sum of the means and the sum of variances of the N independent random variables.

The theorem is even applied when the individual random variables are not Gaussian. In addition, it is applies in certain special cases even when the individual random variables are not independent.

1.8 Random Processes

A random process $X(A, t)$ can be viewed as a function of two variables: an event A and time t. Fig.1.2 illustrates a random process with N sample functions of time, $\{X_j(t)\}$. Each of the sample functions can be regarded as the output of a different noise generator. For a specific event A_j, we have a single time function $X(A_j, t) = X_j(t)$ (i.e., a sample function). The totality of all sample functions is called an ensemble.

For a specific time t_k, $X(A, t_k)$ is a random variable $X(t_k)$ whose value depends on the event. Finally, for a specific event, $A = A_j$ and a specific time $t = t_k$, $X(A_j, t_k)$ is simply a number.

For notational convenience we shall designate the random process by $X(t)$, and let the functional dependence upon A be implicit.

1.8.1 Statistical Averages of a Random Process

Because the value of a random process at any future time is unknown (since the identity of the event A is unknown), a random process whose distribution functions are continuous can be described statistically with a probability density function (***pdf***). In general, the form of the ***pdf*** of a random process will be different for different times. In most situations it is not practical to determine empirically the probability distribution of a random process.

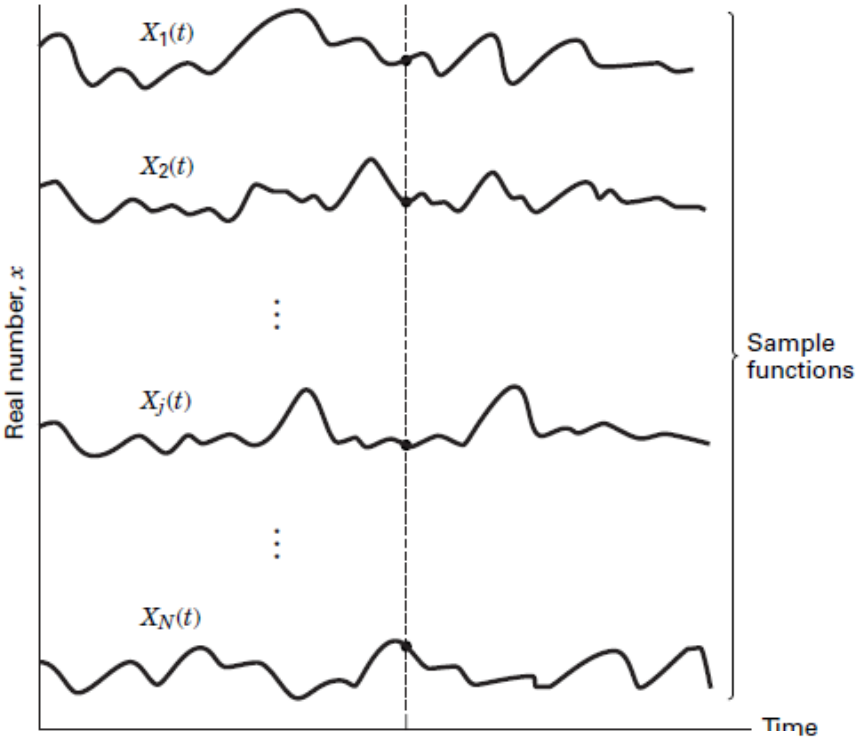


Fig.1.2: Random Noise Process

However, a partial description consisting of the mean and autocorrelation function are often adequate for the needs of communication systems. We define the mean of the random process $X(t)$ as:

$$E\{X(t_k)\} = \int_{-\infty}^{\infty} x\, p_{X_k}(x)\, dx = m_X(t_k) \tag{1.55}$$

where $X(t_k)$ is the random variable obtained by observing the random process at time t_k and the ***pdf*** of $X(t_k)$, the density over the ensemble of events at time t_k, is designated $p_{X_k}(x)$.

We define the autocorrelation function of the random process $X(t)$ to be a function of two variables, t_1 and t_2, given by:

$$R_x(t_1, t_2) = E\{X(t_1)\, X(t_2)\} \tag{1.56}$$

where $X(t_1)$ and $X(t_2)$ are random variables obtained by observing $X(t)$ at times t_1 and t_2, respectively. The autocorrelation function is a measure of the degree to which two time samples of the same random process are related.

1.8.2 Stationarity

A random process $X(t)$ is said to be *stationary* in the *strict sense* if none of its statistics are affected by a shift in the time origin. A random process is said to be *wide-sense stationary* (WSS) if two of its statistics, its mean and autocorrelation function, do not vary with a shift in the time origin. Thus, a process is WSS if

$$E\{X(t)\} = m_x = \text{a constant} \tag{1.57}$$

$$R_X(t_1, t_2) = R_X(t_1 - t_2) \tag{1.58}$$

Strict-sense stationary implies wide-sense stationary, but not vice versa. Most of the useful results in communication theory

are predicated on random information signals and noise being wide-sense stationary. From a practical point of view, it is not necessary for a random process to be stationary for all time but only for some observation interval of interest.

For stationary processes, the autocorrelation function in Equation (1.58) does not depend on time but only on the difference between t_1 and t_2. That is, all pairs of values of $X(t)$ at points in time separated by $\tau = t_1 - t_2$ have the same correlation value. Thus, for stationary systems, we can denote $R_x(t_1, t_2)$ simply as $R_x(\tau)$.

1.8.3 Autocorrelation of Wide-Sense Stationary Process

Just as the variance provides a measure of randomness for random variables, the autocorrelation function provides a similar measure for random processes. For a wide-sense stationary process, the autocorrelation function is only a function of the *time difference* $\tau = t_1 - t_2$; that is,

$$R_X(\tau) = E\{X(t)X(t+\tau)\} \quad \text{for } -\infty < t < \infty \qquad (1.59)$$

For a zero mean WSS process, $R_X(\tau)$ indicates the extent to which the random values of the process separated by τ seconds in time are statistically correlated. In other words, $R_X(\tau)$ gives us an idea of the frequency response that is associated with a random process. If $R_X(\tau)$ changes slowly as τ increases from zero to some value, it indicates that, on average, sample values of $X(t)$ taken at $t = t_1$ and $t = t_1 + \tau$ are nearly the same. Thus, we would expect a frequency domain representation of $X(t)$ to contain a preponderance of low frequencies. On the other hand, if $R_X(\tau)$ decreases rapidly as T is increased. We would expect $X(t)$ to change rapidly with time and thereby contain mostly high frequencies. Properties of the autocorrelation function of a real-valued wide-sense stationary process are as in Table.1.1.

Table.1.1: Properties of Autocorrelation of a Real-valued Wide-sense Stationary Process

n	Property	Meaning
1	$R_X(\tau) = R_X(-\tau)$	Symmetrical in τ about zero
2	$R_X(\tau) \leq R_X(0)$ all τ	maximum value occurs at the origin
3	$R_X(\tau) \leftrightarrow G_X(f)$	Autocorrelation and power spectral density form a Fourier transform pair
4	$R_X(0) = E\{X^2(t)\}$	Value at the origin is equal to the average power of the signal

1.8.4 Time Averaging and Ergodicity

To compute m_X and $R_X(\tau)$ by ensemble averaging, we would have to average across all the sample functions of the process and would need to have complete knowledge of the first- and second-order joint probability density functions. Such knowledge is generally not available.

When the random process belongs to a special class, known as an ergodic process, its time averages equal its ensemble averages, and the statistical properties of the process can be determined by time averaging over a single sample function of the process. For a random process to be ergodic, it must be stationary in the strict sense. An ergodic process is stationary, but, a stationary process is not necessarily ergodic. However, for communication systems, where we are satisfied to meet the conditions of wide-sense stationarity, we are interested only in the mean and autocorrelation functions. We can say that a random process is ergodic in the mean and autocorrelation if:

$$m_X = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)dt \tag{1.60}$$

$$R_X(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)X(t+\tau)dt \tag{1.61}$$

A reasonable assumption in the analysis of most communication signals (in the absence of transient effects) is that the random waveforms are ergodic in the mean and the autocorrelation function. Since time averages equal ensemble averages for ergodic processes, fundan1cntal electrical engineering parameters such as de value, rms value, and average power can be related to the moments of an ergodic random process. Following is a summary of these relationships:

- $m_X = E\{X(t)\}$ is equal to the de level of the signal.
- m_X^2 is equal to the normalized power in the dc component.
- S[nd] moment$E\{X^2(t)\}$, equals total average normalized power
- $\sqrt{E\{X^2(t)\}}$ equals rms value of the voltage or current signal.
- Variance σ_X^2 is the average normalized power in the time-varying or ac component of the signal.
- If the process has zero mean, its variance is the same as the mean square value, or it represents the total poser in the normalized load.
- Standard deviation $\sigma_X$is the rms of ac component of signal.
- If $m_X = 0$, then σ_X is the rms value of the signal.

1.8.5 Power Spectral Density and Autocorrelation of Process

A random process $X(t)$ can generally be classified as a power signal having a power spectral density $G_X(f)$ of the form shown in Equation (1.20). $G_X(f)$ is particularly useful in communication systems, because it describes the distribution of a signal 's power in the frequency domain. The ***psd*** enables us to evaluate the signal power that will pass through a network having known frequency characteristics. We summarize the principal features of ***psd*** functions as follows:

- $G_X(f) \geq 0$ and is always real valued
- $G_X(f) = G_X(-f)$ for $X(t)$ real-valued
- $G_X(f) \leftrightarrow R_X(\tau)$
- $P_X = \int_{-\infty}^{\infty} G_X(f)\, df$ The relationship between average normalized power and ***psd***

1.9 Noise

The term *noise* is used customarily to designate unwanted signals that tend to disturb the transmission and processing of signals in communication systems, and over which we have incomplete control. In practice, we find that there are many potential sources of noise in a communication system. The sources of noise may be external to the system (e.g., atmospheric noise, galactic noise, man-made noise) or internal to the system.

The second category includes an important type of noise that arises from the phenomenon of spontaneous fluctuations of current flow that is experienced in all electrical circuits. In a physical context, the most common examples of the spontaneous fluctuation phenomenon are shot noise, which, as stated in Section 4.10, arises because of the discrete nature of current flow in electronic devices; and thermal noise, which is attributed to the random motion of electrons in a conductor.

However, insofar as the noise analysis of communication systems is concerned, be they analog or digital, the analysis is customarily based on a source of noise called white-noise, which is discussed next.

1.9.1 White Noise

This source of noise is idealized, in that its power spectral density is assumed to be constant and, therefore, independent of

the operating frequency. The adjective "white" is used in the sense that white light contains equal amounts of all frequencies within the visible band of electromagnetic radiation. We may thus make the statement:

> White noise, denoted by $W(t)$, is a stationary process whose power spectral density $S_W(f)$ has a constant value across the entire frequency interval.

Clearly, white-noise can only be meaningful as an abstract mathematical concept; we say so because a constant power spectral density corresponds to an unbounded spectral distribution function and, therefore, infinite average power, which is physically nonrealizable.

Nevertheless, the utility of white-noise is justified in the study of communication theory by virtue of the fact that it is used to model channel noise at the front end of a receiver.

Typically, the receiver includes a filter whose frequency response is essentially zero outside a frequency band of some finite value. Consequently, when white noise is applied to the model of such a receiver, there is no need to describe how the power spectral density $S_{WW}(f)$ falls off outside the usable frequency band of the receiver. Let:

$$S_{WW}(f) = \frac{N_o}{2} \tag{1.62}$$

as illustrated in Fig.1.3.a. Since the autocorrelation function is the inverse Fourier transform of the power spectral density in accordance with the Wiener–Khintchine relations, it follows that for white-noise the autocorrelation function is

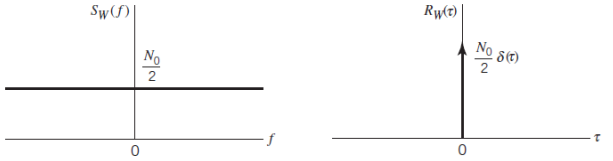
$$R_{WW}(\tau) = \frac{N_o}{2}\,\delta(\tau) \tag{1.63}$$

Hence, the autocorrelation function of white noise consists of a delta function weighted by the factor $N_0/2$ and occurring at the time shift $\tau = 0$, as shown in Fig.1.3.b.

Since $R_{WW}(\tau)$ is zero for $\tau \neq 0$, it follows that any two different samples of white noise are uncorrelated no matter how closely together in time those two samples are taken.

If the white noise is also Gaussian, then the two samples are statistically independent in accordance with Property 4 of the Gaussian process. In a sense, then, white Gaussian noise represents the ultimate in "randomness".

The utility of a white-noise process in the noise analysis of communication systems is parallel to that of an impulse function or delta function in the analysis of linear systems.



(a) Power Spectral Density (b) Autocorrelation Function

(b) Fig.1.3: Characteristics of White Noise

Just as we may observe the effect of an impulse only after it has been passed through a linear system with a finite bandwidth, so it is with white noise whose effect is observed only after passing through a similar system. We may therefore state:

> As long as the bandwidth of a noise process at the input of a system is appreciably larger than the bandwidth of the system itself, then we may model the noise process as white noise.

1.9.2 Ideal Low-pass Filtered White Noise

Suppose that a white Gaussian noise of zero mean and power spectral density $N_0/2$ is applied to an ideal low-pass filter of bandwidth B and passband magnitude response of one. The power spectral density of the noise $N(t)$ appearing at the filter output, as shown in Fig.1.4.a, is therefore:

$$S_{NN}(f) = \begin{cases} \frac{N_0}{2} , -B < f < B \\ 0 \ , \qquad |f| > B \end{cases} \tag{1.64}$$

Since the autocorrelation function is the inverse Fourier transform of the power spectral density, it follows that:

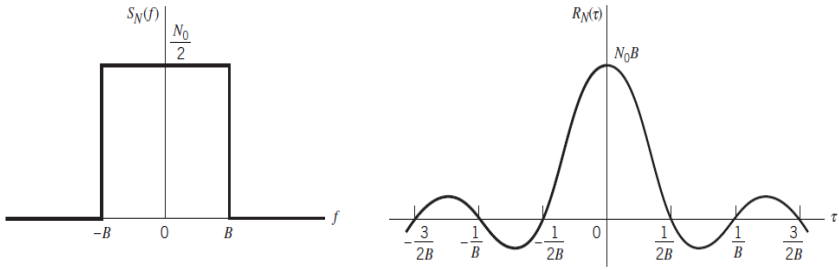
$$R_{NN}(\tau) = \int_{-B}^{B} \frac{N_0}{2} \, e^{j2\pi f\tau} df = N_0 B \, \mathrm{sinc}(2B\tau) \tag{1.65}$$

whose dependence on τ is plotted in Fig.1.4.b. From this figure, we see that $R_{NN}(\tau)$ has the maximum value $N_0 B$ at the origin and it passes through zero at $\tau = \pm k/(2B)$, where $k = 1, 2, 3, \ldots$

Since the input noise $W(t)$ is Gaussian (by hypothesis), it follows that the band-limited noise $N(t)$ at the filter output is also Gaussian. Suppose, then, that $N(t)$ is sampled at the rate of $2B$ times per second.

From Fig.1.4.b, we see that the resulting noise samples are uncorrelated and, being Gaussian, they are statistically independent.

Accordingly, the joint probability density function of a set of noise samples obtained in this way is equal to the product of the individual probability density functions. Note that each such noise sample has a mean of zero and variance of $N_0 B$.



(a) Power Spectral Density (b) Autocorrelation Function

Fig.1.4: Characteristics of Low Pass Filtered White Noise

1.9.3 Correlation of White Noise with Sinusoidal Wave

Consider the sample function:

$$w'(t) = \sqrt{\frac{2}{T}} \int_0^T w(t)\ \cos(2\pi f_c t)\ dt \qquad (1.66)$$

which is the output of a correlator with white Gaussian noise sample function $w(t)$ and sinusoidal wave $\sqrt{2/T}\ \cos(2\pi f_c t)$ as its two inputs; the scaling $\sqrt{2/T}$ factor is included in (4.104) to make the sinusoidal wave input have unit energy over the interval $0 \leq t \leq T$.

With $w(t)$ having zero mean, it immediately follows that the correlator output $\acute{w}(t)$ has zero mean too. The variance of the correlator output is therefore defined by:

$$\begin{aligned}\sigma_{W'}^2 &= E\left[\frac{2}{T}\int_0^T \int_0^T w(t_1)\cos(2\pi f_c t_1) w(t_2)\cos(2\pi f_c t_2)\, dt_1 dt_2\right] \\ &= \frac{2}{T}\int_0^T \int_0^T E[w(t_1)w(t_2)]\cos(2\pi f_c t_1)\cos(2\pi f_c t_2)\, dt_1 dt_2 \\ &= \frac{2}{T}\int_0^T \int_0^T \frac{N_o}{2}\delta(t_1 - t_2)\cos(2\pi f_c t_1)\cos(2\pi f_c t_2)\, dt_1 dt_2 \qquad (4.67)\end{aligned}$$

where, in the last line, we made use of (4.101). We now invoke the *sifting property* of the delta function, namely:

$$\int_{-\infty}^{\infty} g(t)\ \delta(t)\ dt = g(0) \qquad (4.68)$$

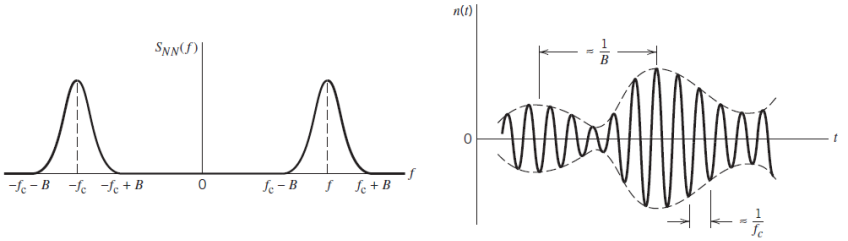
where $g(t)$ is a continuous function of time that has the value $g(0)$ at time t = 0. Hence, we may further simplify the expression for the noise variance as:

$$\sigma_{W'}^2 = \frac{N_o}{2}\frac{2}{T}\int_{-T}^{T}\cos^2(2\pi f_c t)\ dt = \frac{N_o}{2T}\int_0^T [1 + \cos(4\pi f_c t)]dt = \frac{N_o}{2} \qquad (4.69)$$

where, in the last line, it is assumed that the frequency f_c of the sinusoidal wave input is an integer multiple of the reciprocal of T for mathematical convenience.

1.10 Narrowband Noise

The receiver of a communication system usually includes some provision for *preprocessing* the received signal. Typically, the preprocessing takes the form of a *narrowband filter* whose bandwidth is just large enough to pass the modulated component



(a) Power Spectral Density (b) Sample Function

Fig.1.5: Narrowband Noise

of the received signal essentially undistorted, so as to limit the effect of channel noise passing through the receiver. The noise process appearing at the output of such a filter is called *narrowband noise*. With the spectral components of narrowband noise concentrated about some midband frequency $\pm f$c as in Fig.1.5.a, we find that a sample function $n(t)$ of such a process appears somewhat similar to a sine wave of frequency fc. The sample function $n(t)$ may, therefore, undulate slowly in both amplitude and phase, as illustrated in Fig.1.5.b.

Consider, then, the $n(t)$ produced at the output of a narrowband filter in response to the sample function $w(t)$ of a white Gaussian noise process of zero mean and unit power spectral density applied to the filter input; $w(t)$ and $n(t)$ are sample functions of the respective processes $W(t)$ and $N(t)$. Let $H(f)$ denote the transfer function of this filter. Accordingly, we may express the power spectral density $S_N(f)$ of the noise $N(t)$ in terms of $H(f)$:

$$S_N(f) = [H(f)]^2 \tag{1.70}$$

On the basis of this equation, we may now make the following statement:

> *Any narrowband noise encountered in practice may be modeled by applying a white-noise to a suitable filter in the manner described in (4.70).*

In this section we wish to represent the narrowband noise $n(t)$ in terms of its in-phase and quadrature components in a manner similar to that described for a narrowband signal in Section 2.10. The derivation presented here is based on the idea of pre-envelope and related concepts, which were discussed in Chapter 2 on Fourier analysis of signals and systems.

Let $n_+(t)$ and , respectively, denote the pre-envelope and complex envelope of the narrowband noise $n(t)$. We assume that the power spectrum of $n(t)$ is centered about the frequency f_c. Then we may write where is the Hilbert transform of $n(t)$. The complex envelope may itself be expressed as

$$n_+(t) = n(t) + j\hat{n}(t) \tag{1.71}$$

$$\hat{n}(t) = n_+(t)\, e^{-j2\pi f_c t} \tag{1.72}$$

Where $\hat{n}(t)$ is the Hilbert transform of $n(t)$. The complex envelope may itself be expressed as:

$$\hat{n}(t) = n_I(t) + jn_Q(t) \tag{1.73}$$

Hence, combining (4.71) through (4.73), we find that the *in-phase component* $n_I(t)$ and the *quadrature component* $n_Q(t)$ of the narrowband noise $n(t)$ are respectively:

$$n_I(t) = n(t)\cos(2\pi f_c t) + \hat{n}(t)\sin(2\pi f_c t) \tag{1.74}$$

$$n_Q(t) = \hat{n}(t)\cos(2\pi f_c t) - n(t)\sin(2\pi f_c t) \tag{1.75}$$

Eliminating between (4.74) and (4.75), we get the desired *canonical form* for representing the narrowband noise $n(t)$, as shown by:

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \qquad (1.76)$$

Using (4.112) to (4.76), we may now derive some important properties of the in-phase and quadrature components of a narrowband noise, as described next.

First: *The in-phase component $n_I(t)$ and quadrature component $n_Q(t)$ of narrowband noise n(t) have zero mean.*

To prove this property, we first observe that the noise is obtained by passing *n*(*t*) through a linear filter (i.e., Hilbert transformer). Accordingly, $\hat{n}(t)$ will have zero mean because *n*(*t*) has zero mean by virtue of its narrowband nature. Furthermore, from (4.112) and (4.113), we see that $n_I(t)$ and $n_Q(t)$ are weighted sums of *n*(*t*) and $\hat{n}(t)$. It follows, therefore, that the in-phase and quadrature components, $n_I(t)$ and $n_Q(t)$, both have zero mean.

Second: *If the narrowband noise n(t) is Gaussian, then its in-phase component and quadrature component are jointly Gaussian.*

To prove this property, we observe that $\hat{n}(t)$ is derived from *n*(*t*) by a linear filtering operation. Hence, if *n*(*t*) is Gaussian, the Hilbert transform is also Gaussian, and *n*(*t*) and $\hat{n}(t)$ are jointly Gaussian. It follows, therefore, that the in-phase and quadrature components, $n_I(t)$ and $n_Q(t)$, are jointly Gaussian, since they are weighted sums of jointly Gaussian processes.

Third *If the narrowband noise is weakly stationary, then its in-phase component and quadrature component are jointly weakly stationary.*

If *n*(*t*) is weakly stationary, so is its Hilbert transform $\hat{n}(t)$. However, since the in-phase and quadrature components, $n_I(t)$ and $n_Q(t)$, are both weighted sums of *n*(*t*) and $\hat{n}(t)$ and the weighting functions $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$, vary with time, we cannot directly assert that are weakly stationary. To prove Property 3, we have to evaluate their correlation functions.

Using (4.112) and (4.113), we find that the in-phase and quadrature components, $n_I(t)$ and $n_Q(t)$, of a narrowband noise $n(t)$ have the same autocorrelation function, as shown by:

$$R_{N_I N_I}(\tau) = R_{N_Q N_Q}(\tau) = R_{NN}(\tau)\cos(2\pi f_c \tau) + \hat{R}_{NN}(\tau)\sin(2\pi f_c \tau) \quad (1.77)$$

$$R_{N_I N_Q}(\tau) = -R_{N_Q N_I}(\tau) = R_{NN}(\tau)\sin(2\pi f_c \tau) - \hat{R}_{NN}(\tau)\cos(2\pi f_c \tau) \quad (1.78)$$

Where $R_{NN}(\tau)$ is the autocorrelation function of $n(t)$, and $\hat{R}_{NN}(\tau)$ is its Hilbert transform. From (4.77) and (4.78), we readily see that the correlation functions $R_{N_I N_I}(\tau)$, $R_{N_Q N_Q}(\tau)$, and $R_{N_I N_Q}(\tau)$ of the in-phase and quadrature components and depend only on the time shift τ. This dependence, in conjunction with Property 1, proves that $n_I(t)$ and $n_Q(t)$ are weakly stationary if the original narrowband noise is weakly stationary.

Forth: *Both the in-phase noise$n_I(t)$ and quadrature noise $n_Q(t)$ have the same power spectral density, which is related to the power spectral density $S_{NN}(f)$ of the original narrowband noise as follows:*

$$S_{N_I N_I}(f) = S_{N_Q N_Q}(f) = \begin{cases} S_{NN}(f - f_c) + S_{NN}(f + f_c), & -B \le f \le B \\ o, & elsewhere \end{cases} \quad (1.79)$$

Where it is assumed that $S_{NN}(f)$ occupies the frequency interval $f_c - B \le |f| \le f_c + B$ and fc $> B$.

Fifth: *The in-phase and quadrature components $n_I(t)$ and $n_Q(t)$ have the same variance as the narrowband noise $n(t)$.*

Six: *The cross-spectral densities of the in-phase and quadrature components of a narrowband noise are purely imaginary, as shown by:*

$$S_{N_I N_Q}(f) = -S_{N_Q N_I}(f)$$
$$= \begin{cases} j[S_N(f + f_c) - S_N(f - f_c)], & -B \le f \le B \\ 0, & otherwise \end{cases} \quad (1.80)$$

Seventh: *If a narrowband noise $n(t)$ is Gaussian with zero mean and a power spectral density $S_{NN}(f)$ that is locally symmetric about the midband frequency $\pm f_c$, then the in-phase noise $n_I(t)$ and the quadrature noise $n_Q(t)$ are statistically independent.*